

## **Transformation des Wissens und Wandel des Publizierens**

In den Empfehlungen des Wissenschaftsrats zur „Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020“ werden sechs Forschungsformen benannt, die als entscheidende Bezugspunkte für die Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen angeführt werden und deren jeweilige Akteure die Nutzer- und Zielgruppen bilden, auf die sich die Infrastrukturentwicklung ausrichten soll. Bei den sechs Forschungsformen handelt es sich um Forschungsformen, die experimentierend, simulationsbezogen, beobachtend, hermeneutisch-interpretierend, begrifflich-theoretisch oder gestaltend sind<sup>1</sup>. Dabei wird ausdrücklich darauf hingewiesen, dass die genannten Forschungsformen mit unterschiedlicher Ausprägung und Intensität in den einzelnen Wissenschaftsdisziplinen praktiziert werden, dass innerhalb eines Faches aber durchaus auch mehrere Forschungsformen zum Einsatz kommen. Als Beispiele dafür werden Sprachwissenschaft und Physik genannt, wo beobachtende, experimentierende, begrifflich-theoretische sowie simulationsbezogene Forschungsformen mit unterschiedlicher Schwerpunktsetzung parallel oder supplementär eingesetzt werden.

Mit dieser Typologie der Forschungsformen wird die herkömmliche Differenzierung der Wissenschaftsdisziplinen in Geistes-, Sozial-, Natur-, Lebens- und Technikwissenschaften von einer Systematik abgelöst, die stärker methodische, als epistemische Ansätze in den Blick nimmt und auf diese Weise die jeweiligen Forschungsformen mit den jeweils genutzten Medien „als Objekt, Quelle und Speicher des Erkenntnisgewinns“ verbindet. Insofern stellen die jeweils genutzten Medien die Bezüge der Forschungsformen zur Informationsinfrastruktur her, wie umgekehrt die Medien die Bezüge der Informationsinfrastruktur zu den Forschungsformen herstellen. Dieser Umkehrschluss wird in den WR-Empfehlungen zwar nicht unmittelbar vorgenommen, da sich die Ausrichtung der Informationsinfrastruktur ausdrücklich an der jeweiligen „Forschungsform“ ori-

1 Vgl. Empfehlungen des Wissenschaftsrats zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020 (2012) und <http://www.wissenschaftsrat.de/download/arciv/23259-12.pdf> S. 35ff.

entieren soll. Von daher lassen sich die Empfehlungen zur Infrastrukturentwicklung sowohl auf digitale als auch auf nicht-digitale, analoge Informationsstrukturen beziehen, wie mehrfach mit Nachdruck betont wird. Doch mit den skizzierten Potenzialen digitaler, vernetzter Forschungsaktivitäten, die bei „nahezu allen Forschungsformen“ zu Erleichterung und Erweiterung der wissenschaftlichen Arbeit führen und mit denen sich neue Forschungsfragen und Forschungsgebiete sowie neue „Perspektiven auf bereits bearbeitete Forschungsgegenstände“ auftun, wird deutlich, dass es „eine Vielzahl von Rückkopplungen zwischen den Wissensordnungen der Wissenschaft und den von ihr genutzten Informationsinfrastrukturen gibt“, wie es in diesem Kontext der Empfehlungen heißt. Diese Formulierung gibt ganz unmissverständlich zu erkennen, dass zwischen den Forschungsformen, den Medien und dem Spektrum der Informationsinfrastruktur ein wechselseitiger Einfluss gesehen und für die Weiterentwicklung als maßgeblich betrachtet wird.

In den kurzen Beschreibungen zu den Forschungsformen kommt den jeweils genutzten Medien besondere Bedeutung zu. Damit wird verdeutlicht, welche Medien für welche Forschungsformen charakteristisch und prägend sind. Bemerkenswert ist dabei, dass mit Ausnahme der hermeneutisch-interpretierenden Forschungsformen für das gesamte Spektrum der Forschungsformen grundsätzlich Forschungsdaten als mediale Voraussetzung für die wissenschaftliche Arbeit der Ausgangspunkt sind. Angesichts dessen lässt sich mit Blick auf die Forschungsformen von „data driven research“ sprechen, wie es insbesondere für die Lebens-, Natur- und Technikwissenschaften kennzeichnend ist. Die für die hermeneutisch-interpretierenden Forschungsformen primär vermutete Nutzung analoger Medien (Artefakte, Bilder, Bücher, Objekte, Schriften etc.) wird insofern etwas zurückgenommen, als diese nicht-digitalen Materialien auch in digitalisierter Form sowie als „digitally born“ Materialien zur Verfügung stehen können. Ohne dass dies im Text der Empfehlungen ausdrücklich angesprochen wird, schließt sich damit der Kreis der von Daten getriebenen Wissenschaft für alle Forschungsformen. Denn wenngleich sich digitalisierte oder „digitally born“ Dokumente oder Objekte, die hermeneutisch-interpretierenden Forschungsformen zugrunde liegen, von experimentellen, simulationsbezogenen oder beobachteten Forschungsdaten wesentlich unterscheiden, handelt es sich dennoch um Daten bzw. um Datenbestände. In Anbetracht dessen wird mit den genannten Forschungsformen aus verschiedenen Perspektiven ein Daten-Universum skizziert, das aufgrund der digitalen Transformation für alle Wissenschaftsbereiche bestimmend ist und zu neuen Informationsinfrastrukturen führt. Für die Ausgestaltung und Weiterentwicklung der Informationsinfrastrukturen ist deshalb das Paradigma des „data driven research“ maßgeblich. Zugleich hat damit das Daten-Paradigma

das Selbstverständnis der Wissenschaft sowie das Verständnis von Wissen und Wissensgütern erreicht. Ist damit „künstliche Intelligenz“ und „datengetriebener Empirismus“ für e-research und e-science leitend?

### *1. Digitalisierung der Informationsinfrastrukturen*

Die Digitalisierung im Alltag unterschiedlicher Lebenswelten hat in Verbindung mit einer intensiven Nutzung des Internets zu einer massiven Überflutung der Verbrauchermärkte mit Endgeräten, mobilen Devices, Peripheriegeräten, Apps, Tools und einer Vielzahl von Software-Komponenten geführt. Die sich daraus ergebenden Innovationen liegen schwerpunktmäßig in der explosionsartigen Generierung, Verbreitung und Rezeption beliebiger Informationen und Inhalte sowie in der signifikant expandierenden, digitalen Kommunikation, die weltweit nahezu alle Ebenen des gesellschaftlichen Lebens durchdringt. In Forschung und Lehre hat sich mit der Digitalisierung ebenfalls ein Paradigmenwechsel vollzogen. Dazu gehören die Nutzung von Computer und Internet für die wissenschaftliche Arbeit sowie im Kontext der Bürokommunikation, die Digitalisierung der Wissenschaftskommunikation mit elektronischen Büchern und Zeitschriften unter Einschluss von retro-digitalisierten Materialien sowie die zahlreichen Applikationen – oder auch Geräte - zur Unterstützung von Arbeitsabläufen und Prozessen in Forschung, Lehre, Studium und Verwaltung; in diesen Zusammenhang gehören auch die IT-gestützten Verwaltungssysteme der Bibliotheken. War der Wissenschaftsbereich in den 80er und 90er Jahren des 20. Jahrhunderts zunächst Impulsgeber und Treiber für den Einsatz von IT-Anwendungen auch über Hochschulen und Forschungseinrichtungen hinaus, so hat sich diese Entwicklung nicht fortgesetzt, sondern wurde von der Marktentwicklung einer boomenden IT-Industrie eingeholt. Die Anbieter von mobilen Devices (smart technologies), Kommunikationsplattformen (social networks) und mächtigen Suchmaschinen einschließlich intuitiv nutzbarer Tools (app-galleries) haben die Entwicklungsführerschaft übernommen und bedienen mit hoher Dynamik einen weltweit präsenten Markt, an dem zahlreiche Content- und Service-Provider unterschiedlicher Größenordnung partizipieren. Dabei ist entscheidend, dass die jeweils eingenommenen Marktpositionen Gewinn- und Umsatzpotenziale bieten, die den Firmen die wirtschaftliche Überlebensfähigkeit und den Marktanschluss trotz hoher Entwicklungsaufwände und kundenfreundlicher Preise sichern.

Die damit einher gehende Marktentwicklung mag verdeutlichen, wie sich die Ausrichtung der ursprünglich durchaus entwicklungsorientierten „Hochschul-IT“ von den Profilen der nun Markt beherrschenden Player unterscheidet. Anders gesagt: Die Weiterentwicklung der digitalen Transformation findet heute

verstärkt außerhalb des Wissenschaftsbereichs statt, obwohl diese Entwicklung ursprünglich von Forschungseinrichtungen und Hochschulen ihren Ausgang nahm. Zugleich wird dadurch der Eindruck vermittelt, dass der vielfach diagnostizierte Innovationsdruck „von außen“ Forschung und Lehre nicht erreicht und insofern bisher nicht zu den notwendigen Veränderungen in der „Academia“ führte, an der sich Vorreiter von Innovation oder „early adopter“ eher in der Minderheit sehen und bisweilen sogar isoliert erscheinen. Etwas anders sieht dies bei zentralen Serviceeinrichtungen wie Bibliotheken, Medien- und Rechenzentren aus, die sich als Informations- oder IT-Dienstleister in der Situation befinden, die Herausforderungen der Digitalisierung aufzugreifen und mit der Entwicklung entsprechender Dienstleistungen umzusetzen. Dabei wird allerdings immer wieder deutlich, dass für die Akzeptanz innovativer Informations- und Mediendienste viel Überzeugungsarbeit erforderlich ist.

Ganz unbestritten ist, dass der Wissenschaftsbereich andere Anforderungen an den Einsatz innovativer Technologie hat als die jeweils aktuelle und sich permanent wandelnde „Alltags-IT“, die schon aus fachlichen Gründen nicht unmittelbar auf den Wissenschaftsbereich übertragbar ist. Zugleich sind die Einrichtungen der Informationsinfrastruktur an Hochschulen aus finanziellen oder kapazitären Gründen oft nur eingeschränkt in der Lage, die Weiterentwicklung ihrer Serviceportfolios an die Entwicklungsdynamik marktgetriebener Innovationen anzupassen. Dabei sind über fehlende Ressourcen hinaus insbesondere rechtliche Rahmenbedingungen zu nennen wie beispielsweise die Einhaltung datenschutzrechtlicher Regelungen oder die Gewährleistung von Datensicherheit. Diese Rahmenbedingungen führen oft dazu, dass hochschul- bzw. wissenschaftseigene Angebote unflexibel und zu wenig komfortabel erscheinen, um wirklich „gerne“ genutzt zu werden. Verglichen mit den Angeboten auf dem Verbrauchermarkt kommen schließlich viele Entwicklungen wissenschaftseigener Dienste und Infrastrukturen zu spät. Dies wiederum führt dazu, dass die im Regelfall einfach nutzbaren, oft sehr preiswerten Angebote großer IT-Provider wie insbesondere Kollaborationstools, Mail, Online-Speicher und „social networks“ doch in den Wissenschaftsbereich Eingang finden, ohne dass die im Grundsatz erforderlichen Voraussetzungen des Datenschutzes, der Datensicherheit und der Datenhoheit gegeben sind. In Anbetracht dessen ist und bleibt der Aufbau wissenschaftsspezifischer Infrastrukturen auch künftig ein dringendes Desiderat.

Vor diesem Hintergrund hat es in den vergangenen 20 Jahren weltweit auf internationaler und nationaler Ebene sowie innerhalb und außerhalb des Wissenschaftsbereichs eine Vielzahl an Empfehlungen, Förderinitiativen und Projekte gegeben, die die Herausforderungen der digitalen Transformation für Forschung, Lehre, Studium und Verwaltung aufgegriffen und in unterschiedlichem Umfang

zur Realisierung gebracht haben. Die Stichwörter dieser Entwicklung sind für die Forschung „e-research“ und „e-science“, für Lehre und Studium „e-learning“ und „e-teaching“ und für den Bereich der Verwaltung „e-administration“ und „e-governance“. Dass die Vorsilbe „e-“ eingeführt wurde, um die „digitale Hochschule“ von der „analogen Hochschule“ zu unterscheiden und die herkömmlichen Verfahren von den neuen, technologiegetriebenen Verfahren im Sinne von „enhanced“ abzusetzen, soll hier ebenso wenig vertieft werden wie der immer wieder neue Versuch, die Veränderungspotenziale der Informationstechnologie für die Einführung unternehmerischer Strukturen im Sinne von „business models“ in Forschung, Lehre und Studium einzusetzen. Was vor allem mit der Hochschule als unternehmerischem Geschäftsmodell adressiert wird, ist ein Kulturwandel, der im Zuge der Digitalisierung des akademischen „Kerngeschäfts“ möglich erscheint und dabei zu einem veränderten, neuen Selbstverständnis der „academia“ führen kann. Im vorliegenden Zusammenhang soll dagegen der mit der digitalen Transformation einhergehende Wandel des Wissensverständnisses betrachtet werden, der im direkten Zusammenhang mit der Weiterentwicklung der Informationsinfrastrukturen steht. Denn bei den Aktivitäten und Initiativen, die unter dem Label „enhanced education and enhanced research“ anvisiert und umgesetzt werden, ist eine unmittelbare Verbindung mit der jeweils erforderlichen Informationsinfrastruktur gegeben, die zwingende Voraussetzung für die digitale Transformation von Forschung, Lehre und Studium ist.

Vor diesem Hintergrund bestehen digitale Informationsinfrastrukturen zum einen im Sinne einer eher informationszentrierten Unterstützung des wissenschaftlichen Arbeitens. Dabei geht es um Verfügbarkeit und Verbreitung von Inhalten, die selbst in digitaler Form vorliegen oder als digitale Objekte generiert werden. Die dafür notwendige Unterstützung wird von Bibliotheken angeboten, die e-books und e-journals lizenzieren, Recherche- und Beratungsservices anbieten, Forschungsergebnisse nach den Prinzipien des Open Access publizieren, analog verfügbare Materialien im Rahmen der Möglichkeiten der Urheberrechtsgesetzgebung (retro-)digitalisieren und die sich um die langfristige Verfügbarkeit bzw. Archivierung digitaler Inhalte (Bilder, Daten, Texte etc.) bemühen. Zum anderen - und darüber hinaus gehend - sind künftig verstärkt von Bibliotheken anzubietende Infrastrukturdienste zu nennen, die kollaborative Arbeitsformen und die Aufbereitung und Verarbeitung von Inhalten durch Annotieren, Kontextualisierung, Referenzierung und Strukturierung sowie Data- und Textmining oder andere Retrievalformen unterstützen. Dabei geht es im Unterschied zu den zuvor genannten Unterstützungsmaßnahmen um Dienste zur informationstechnischen Unterstützung, die sich unmittelbar auf das wissenschaftliche Arbeiten beziehen und damit fester Bestandteil von Forschungsprozessen sind. Dabei wird der epis-

temische Zusammenhang zwischen Infrastruktur und Wissensgenerierung deutlich, der im Grundsatz nicht neu ist, aber in digitaler Form eine neue Qualität enthält. Denn maschinelle Arbeits- und Auswertungsverfahren ersetzen analoge oder manuelle Prozesse und enthalten ein ungleich größeres Potenzial an Effektivität und Effizienz, als dies mit herkömmlichen Verfahren der Verarbeitung und Evaluation möglich und vorstellbar ist.

## *2. Wissen im Wandel*

Wenn IT-gestützte Informationsinfrastrukturen eng mit der Produktion, der Repräsentation und der Rezeption von Wissen und Wissensgütern in Zusammenhang stehen, werden die Prozesse der wissenschaftlichen Arbeit verstärkt automatisiert und damit maschinellen Verfahren unterworfen. Solche Verfahren sind in der Geschichte des Wissens nicht ungewöhnlich und geschehen mit der Digitalisierung gewiss nicht zum ersten Mal. Denn schon immer haben Menschen sich darum bemüht, große Herausforderungen an Technik auszulagern. Ein sehr gutes Beispiel dafür ist das Automobil, an das vor allem körperliche Bewegung von Stadt A nach Stadt B ausgelagert wird. Seit vielen Jahrhunderten – um nicht zu sagen, seit einigen Jahrtausenden – lagern wir unser Gedächtnis zunächst an die Technik der Schrift, dann an beschriftete Trägermedien (Papyrus, Pergament, Papier) und nun an Maschinen, Computer und Netze, aus. Indem wir Verarbeitung (Processing), Verbreitung (Distribution) und Speicher (Memory) auf Server und Netze auslagern können, sind wir in der Situation, in einem Umfang auszulagern, der uns bisher nicht zur Verfügung stand. Wir konnten noch nie so viel an Gedächtnis, Verarbeitung und Verbreitung auslagern wie es uns heute möglich ist<sup>2</sup>. Aber wir laufen zugleich Gefahr, durch Auslagerung auch zu verlieren, so dass sich die Frage stellt: Was machen wir mit diesen neuen Maschinen und was machen diese Maschinen mit uns? Mit dieser Frage wird deutlich, dass die digitale Transformation mit Entwicklungen der künstlichen Intelligenz („artificial intelligence“) in Zusammenhang steht, in der mit informationstechnischen Ansätzen versucht wird, kreative Intelligenz zu modellieren und mit Hilfe technischer Verfahren zu automatisieren. Mag ein in diesem Sinne vollständiger Ansatz die Möglichkeiten der künstlichen Intelligenz überschreiten, so zählen einzelne Komponenten wie Information Retrieval, Suchmaschinen,

2 Vgl. Degkwitz, A., Texte, Daten, Bilder – Wissen! – In: Bibliothek, Forschung und Praxis 36 (2012) 3, S. 215 – 219. und Zimmerli, W., Vom Unterschied, der einen Unterschied macht. Information, Netzwerkdenken und Mensch-Maschine-Tandem. – In: Die Zukunft des Wissens. XVIII. Deutscher Kongress für Philosophie (1999). Hrsg. v. Jürgen Mittelstrass. - Berlin, Akademie Verlag, 2000, S. 120 – 122.

Sprach- und Zeichenerkennung, Text- und Datamining, maschinelles Übersetzen oder wissensbasierte Systeme zu den Anwendungsfeldern, die als Teilgebiete der künstlichen Intelligenz beforscht wurden und werden<sup>3</sup>.

Unter epistemischen Gesichtspunkten wird im Zusammenhang mit der Digitalisierung wissenschaftlicher Arbeitsformen von einer „neuen“ Qualität des Wissens gesprochen, indem mit den Möglichkeiten digitaler Verarbeitung und Vernetzung von Daten und Texten – nicht zuletzt auch für Fragen und Themen, die bereits als erforscht gelten – neue Erkenntnisziele identifiziert und damit neue Forschungsergebnisse gewonnen werden können, die aufgrund der Vollständigkeit der jeweils zugrunde liegenden Datenbasis und der darauf aufsetzenden Werkzeuge mit analogen Methoden nicht hätten erreicht werden können. Erkenntnisziele und Wissensergebnisse stellen sich in Geistes- und Naturwissenschaften, verbunden mit den eingangs benannten Forschungsformen, häufig ganz unterschiedlich dar. In den Naturwissenschaften, die von beobachtenden, experimentierenden oder simulationsbezogenen Forschungsformen stark geprägt sind, bilden Daten in nahezu allen Formen die Grundlage wissenschaftlicher Arbeit. Von daher werden seit langem Daten und Datenbestände mit Hilfe der „computational science“ analysiert, evaluiert, simuliert und visualisiert, um anhand entsprechender Modelle, die an beobachteten Phänomenen orientiert sind, Kausalitäten und Kontexte zu erklären und um Gesetzmäßigkeiten abzuleiten oder Entwicklungen zu prognostizieren<sup>4</sup>. Von daher wird in den Naturwissenschaften vorrangig mit Ansätzen der Statistik und Wahrscheinlichkeitsverteilung gearbeitet. Im Unterschied zu naturwissenschaftlichen Forschungen stehen für Geistes-, Kultur- und Sozialwissenschaften primär das Verstehen von menschlichem Denken, Handeln, Schaffen und Wirken im Mittelpunkt. Vor diesem Hintergrund stellt sich als Frage, ob diese auf Daten beruhenden Ansätze auch auf hermeneutisch-interpretierende Forschungsformen übertragbar sind und sich in den „digital humanities“ mit diesen Methoden „Verstehenszusammenhänge“ sinnvoll erforschen lassen. Denn wenn datenbasierte Arbeitsverfahren in den „humanities“ und in den „social sciences“ eingesetzt werden, steht zu erwarten, dass die hermeneutisch-interpretierenden Forschungsformen, die für diese Wissensgebiete einen hohen Stellenwert haben, absehbar durch empirische Metho-

3 Die deutsche Sprache im digitalen Zeitalter – The German Language in the Digital Age. Meta-Net – Weißbuch-Serie. Berlin – Heidelberg: Springer 2012, 81 S.

4 Siehe Zimmerli, W., Wissenskulturen des 18. und 21. Jahrhunderts. – In: Kulturen des Wissens im 18. Jahrhundert. Hrsg. Von Johannes Schneider. – Berlin: de Gruyter, 2008, S. 20 ff.: – Zitat (S. 22): „Galt bislang Georg Simmels Bestimmung, gebildet sei, wer wisse, wo er finde, was er nicht weiß, gilt nun: Gebildet ist, wer weiß, wie er machen kann, was er nicht weiß.“

den und quantitative Verfahren ergänzt und möglicher Weise sogar durch diese ersetzt werden<sup>5</sup>.

Anknüpfend an die herkömmlichen, analogen Verfahren der hermeneutisch-interpretierenden Forschungsformen werden auch in den „digital humanities“ in einem umfassenden Sinne „Artefakte“ wie insbesondere Bilder und Texte zunächst gesammelt und ausgewertet (analysiert). Im digitalen Kontext stellen die technische Aufbereitung sowie die XML-Strukturierung der „Artefakte“ für maschinelle Analyse und Weiterverarbeitung eine wesentliche Voraussetzung dar. Gegenüber analogen Arbeitsweisen sind Aufbereitung und Strukturierung im digitalen Zusammenhang ein zusätzlicher und neuer Teilprozess. Im nächsten Schritt werden die Beschreibungen und Analyseergebnisse des Bild- und Textmaterials referenziert und semantisch in Beziehung gesetzt (kontextualisiert). Schließlich werden die Ergebnisse der Analyse-, Evaluations- und Kontextualisierungsverfahren aufbereitet und veranschaulicht (visualisiert), um anhand der für das Forschungsvorhaben herangezogenen Artefakte die erkenntnisleitenden Fragestellungen zu menschlichem Verhalten und schöpferischem Wirken zu beantworten und in ihrer Kausalität zu verstehen. Datenbasierte Analysen, die quantitative Auswertungen, semantische Kontextualisierung und die Visualisierung von Ergebnissen ermöglichen, können dabei ganz wesentlich zu neuen Erkenntnissen beitragen wie auch zu neuen Forschungsfragen führen<sup>6</sup>. Die Herausforderungen, die sich dabei stellen, sind darin zu sehen, dass die „Quelldaten“ der Geistes-, Kultur- und Sozialwissenschaften überwiegend nicht digital vorliegen, sondern zunächst zu digitalen Daten transformiert werden müssen, um in der weiteren Folge als Datenmaterial verarbeitet werden zu können. Dies gilt im Übrigen auch für digitale „Quelldaten“, die nicht in verarbeitungsfähigen Formaten zur Verfügung stehen. Mit anderen Worten geht es darum, dass das „Implizite“ der zunächst nicht interoperablen bzw. nicht verarbeitbaren „Quelldaten“ geöffnet und „expliziert“ wird. Als Wissenschaften des Diskurses und Verstehens sind den Geisteswissenschaften damit eine Verbreiterung der bisher zugänglichen Datengrundlage und „eine Explizierung des in unseren Daten- und Wissensspeichern inhärenten Wissens verfügbar“, was – „verbunden mit einer ge-

- 5 Siehe Blanke, T. / Horstmann, A., Bericht über das Symposium zum zehnjährigen Bestehen der Abteilung Forschung und Entwicklung. – In: Evolution der Informationsinfrastruktur. Kooperation zwischen Bibliothek und Wissenschaft. Erschienen im Rahmen des zehnjährigen Jubiläums der Abteilung Forschung und Entwicklung der SUB Göttingen, Hrsg. v. H. Neuroth, N. Lossau, A. Rapp. Glückstadt: Verlag Werner Hülsbusch 2013, S. 55ff.
- 6 Vgl. Görtz, G., KI und Geisteswissenschaften. – In: KI – Künstliche Intelligenz. Organ des Fachbereichs „Künstliche Intelligenz“ der Gesellschaft für Informatik e. V. – German Journal of Artificial Intelligence 25 (2011) 4, S. 313 – 315.



wissen Annäherung an naturwissenschaftliche Zugänge“ – für ihre Forschungsfragen Gewinn bringend ist, ohne dabei etwas aufgeben zu müssen. Von daher verbinden sich mit dem Zugriff auf bisher nicht zugängliche Quellen oder einfach nur „mehr“ Quellen für die Geisteswissenschaften neue und eben „mehr“ Erkenntnisse<sup>7</sup>.

Zugleich stellt sich die Frage der Leistungsfähigkeit und der Transparenz von Algorithmen und Werkzeugen, die zur Verarbeitung der dann digital transformierten Daten eingesetzt werden<sup>8</sup>. Sind maschinelle Verfahren zur Analyse, Auswertung und Kontextualisierung für interpretierende Forschungsformen geeignet und angebracht? Ist uns bewusst, was auf den Maschinen passiert, auf die wir unser „Gedächtnis“ und die Verarbeitung unserer „Erinnerung“ auslagern? Möglicher Weise ist dabei nicht so sehr der Einsatz maschineller Verfahren kritisch, sondern vielmehr ein „blindes“ und weitgehend kenntnisfreies Vertrauen auf eine hohe Leistungsfähigkeit und Objektivität der eingesetzten Verfahren, mit denen Daten maschinell generiert oder maschinell generierte Daten genutzt werden, die dann die Grundlage für Analysen und Auswertungen sind. Denn auf diese Weise wird eine Realität geschaffen, die keine konkrete, physische Realität, keine spekulativ gebrochene Realität, sondern eine virtuelle Realität ist – eine Realität, die maschinell simuliert ist und von daher in ihrem Wesen und Wirken der physischen Realität gleichgesetzt wird und dabei auf Daten und Datenbeständen beruht, die die Grundlage für die Verarbeitung und Aggregation empirischen Wissens sind. Anders gesagt: Wir wissen, indem wir generieren, was wir nicht wissen und kommen auf diese Weise zu Ergebnissen, die wir als Empirie verstehen. Dass in dieser virtuellen Umgebung auch die Gefahr besteht, Wirklichkeitsbezüge nur eingeschränkt herzustellen oder gar zu verlieren, ist der Preis der Auslagerung unseres „Gedächtnisses“ an Maschinen.

Schließlich bieten uns die virtuellen Arbeitsumgebungen gewaltige Mehrwerte, indem sie eine Mehrfachverarbeitung von Datenbeständen erlauben, die uns, wie schon gesagt, neue Forschungsfragen und Wissensgebiete eröffnen. Alle Da-

7 Siehe Rapp, A., Aus Sicht der Geisteswissenschaften: Die wissenschaftliche Bibliothek als Schnittstelle zwischen digitaler Infrastruktur und geisteswissenschaftlicher Forschung. – In: Evolution der Informationsinfrastruktur. Kooperation zwischen Bibliothek und Wissenschaft. Erschienen im Rahmen des zehnjährigen Jubiläums der Abteilung Forschung und Entwicklung der SUB Göttingen. Hrsg. v. H. Neuroth, N. Lossau, A. Rapp. Glückstadt: Verlag Werner Hülsbusch 2013, S. 347.

8 Siehe Enke, H. / Fritzsche, B., Erfahrungen aus dem Projekt „WissGrid – Grid für die Wissenschaft“. – In: Evolution der Informationsinfrastruktur. Kooperation zwischen Bibliothek und Wissenschaft. Erschienen im Rahmen des zehnjährigen Jubiläums der Abteilung Forschung und Entwicklung der SUB Göttingen. Hrsg. v. H. Neuroth, N. Lossau, A. Rapp. Glückstadt: Verlag Werner Hülsbusch. 2013, S. 186.

ten, die in dieser Weise quantitativ verarbeitet oder semantisch kontextualisiert werden, können zu neuen Erkenntnissen oder zu neuen Fragestellungen führen, die weitere Verarbeitungsprozesse veranlassen und neue Forschungsprojekte - auch für hermeneutisch-interpretierende Fachgebiete - initiieren. Für die Produktion, Transparenz und Verbreitung von Wissen und Wissensgütern ist das ein enormer Gewinn<sup>9</sup>. Die jeweiligen Eigenheiten und paradigmatischen Unterschiede der Geistes- und Naturwissenschaften werden dabei weder ausgeräumt noch gleichgeschaltet, sondern bleiben bestehen. Doch bieten sich neue Möglichkeiten, Methoden und Verfahren wechselseitig zu übernehmen, weil nun auch in den Geisteswissenschaften Daten und Datenbestände die Grundlage der wissenschaftlichen Arbeit sind.

### *3. Forschungsprozess und Publikation*

Oft wird der Wandel von analogen (gedruckten) zu digitalen (elektronischen) Medien als unmittelbarer Auslöser für die digitale, vernetzte Transformation von Informationen und Kommunikationsprozessen betrachtet. Zugleich unterliegt der Medienwandel dem Einfluss von IT-gestützten Arbeitsverfahren und Workflows, die entweder aus Gründen der Optimierung bestehender Prozesse oder für die Herstellung neuer Produkte eingesetzt werden, die sich mit herkömmlichen Verfahren nicht produzieren lassen. Für das wissenschaftliche Publizieren steht die Digitalisierung der Medien in direktem Zusammenhang mit dem Forschungsprozess, dessen Verlauf immer wieder zu Ergebnispräsentationen und damit zu Publikationen führt, die zugleich als mediale Präsentationen zu sehen sind. Denn Wissen und Medien hängen eng miteinander zusammen und beeinflussen sich wechselseitig: Ohne mediale Transformation kann Wissen weder vermittelt noch verbreitet werden. Medien repräsentieren unser Wissensverständnis. Wissensparadigmen prägen die Gestaltung der medialen Präsentation. Die mediale Gestaltung von Wissen beruht deshalb nicht nur auf Technologie, sie ist auch in den unterschiedlichen, disziplinbezogenen Wissenskulturen begründet.

Über viele Jahrhunderte wurde im europäischen Kontext Wissen in den unterschiedlichen Formaten gedruckter Texte zusammengeführt und veröffentlicht. Ausschließlich textbasiertes Wissen und seine gedruckten Präsentationen verlieren jedoch an Bedeutung. Stattdessen entstehen verstärkt datengestützte Publikationsformen, die multimedial angereichert sind und sich in beliebige Kollaborationskontexte einbringen lassen oder auf kontextualisierten Daten unterschiedlicher Provenienz basieren. In diesem Zusammenhang gehören beispielsweise

9 Siehe Blanke, T. / Horstmann, A., a.a.O. S. 63f. und Rapp, A., a.a.O. S. 349.

Druckbestände, die im OCR-Verfahren als Volltexte digitalisiert sind und für ihre weitere Verarbeitung nach XML strukturiert werden. Darüber hinaus bietet die Nutzung von Softwarekomponenten aus App-Galerien oder App-Stores die Möglichkeit, datenbasiertes Wissen unter verschiedenen Aspekten und Fragestellungen zu verarbeiten. Unterschiedliche Informationsmaterialien können in einem digitalen Dokument zusammen geführt werden und damit Bestandteile einer weiter verarbeitbaren Publikation bilden oder sie werden als differenziert verlinkte Datenbestände veröffentlicht.

Die neuen, digitalen Medienformate übersteigen damit das Potenzial gedruckter Formate signifikant. Dies hat Auswirkungen auf Vermittlung und Verbreitung künftigen Wissens. Hinzu kommt das außerordentlich hohe Potenzial an Kommunikationsmöglichkeiten des Internets, das als Informationsreservoir weltweit just in time zur Verfügung steht. Die „single channel communication“ gedruckter Publikationen wird zunehmend durch die „networked communication“ digitaler Medien ersetzt. Der tief greifende Wandel hat das Potenzial, Dokumente und Texte nun zu multimedialen Präsentationen zu erweitern und den linearen ‚step by step process‘ der gedruckten Veröffentlichungspraxis durch die ‚all to all communication‘ des Internets als primärer Verbreitungsform für wissenschaftliche Publikationen zu ersetzen. Dies lenkt den Blick auf den Forschungsprozess, der auch in den Geistes und Kulturwissenschaften als genuin IT-gestützter Prozess nicht mehr ausschließlich textbasiert, sondern zunehmend auf der Grundlage von Daten erfolgt.

Grundlegende Phasen des Forschungsprozesses aller Fachdisziplinen und Forschungsformen sind vor dem Hintergrund der jeweiligen Forschungsfragen (1) die Generierung oder Sammlung von Materialien, (2) deren Aufbereitung, Strukturierung und Aggregation, (3) Analyse und Evaluation der Materialien und (4) die Ableitung und Gewinnung von Ergebnissen und deren Publikation. Wenn diesen Phasen des Forschungsprozesses Daten und Datenbestände als Materialien zugrunde gelegt werden, ergeben sich daraus Arbeitsvorgänge und Verfahren, die den Forschungsprozess bis zur Publikation unterstützen. Dabei sollen und werden verarbeitete Daten zugleich auch Bestandteil der Publikation sein und die Veröffentlichung der Ergebnispräsentation prägen. Für Generierung oder Sammlung von Materialien werden Daten entweder produziert oder gewonnen oder mit Hilfe von Suchwerkzeugen identifiziert. Zur Aufbereitung und Aggregation von Materialien müssen Daten beschrieben, referenziert und strukturiert werden. Für Analyse und Evaluation werden Retrievalverfahren wie beispielsweise Data- und Textmining eingesetzt. Zur Ergebnisaufbereitung und Veröffentlichung werden über textuelle Ergebnispräsentationen hinaus Visualisierungen und multimediale Anreicherungen genutzt. Hinzu kommen die vielfältigen Möglichkeiten,

Publikationen zu kommentieren, zu verlinken und weiterzuverarbeiten. Da es bei datengetriebenen Forschungsprozessen und datenbasierten Ergebnisveröffentlichungen um die Speicherung und Verarbeitung großer Datenmengen („big data“) geht, sind hierfür virtualisierte und hochgradig skalierbare Systemumgebungen („clouds“) unerlässlich. Für digitale Publikationen, die sich vor diesem Hintergrund als „Enhanced Publications“ verstehen, lassen sich die folgenden Kriterien identifizieren.<sup>10</sup>

(1) „Enhanced Publications“ sind angereichert oder „enriched“. Das heißt, dass sie in beliebigem Umfang Bilder, Fotos, Audios, Videos, digitalisierte Texte, Forschungsdaten, Animationen und Simulationen umfassen oder verlinken. Die Möglichkeit, unterschiedliche Medien und Medienformate in elektronischer Form in Publikationen zu integrieren, ist bei digitalen Publikationen deutlich größer als bei Druckwerken und stellt insofern ein Spezifikum von „Enhanced Publications“ dar.

(2) Enhanced Publications sind „social“. Die Möglichkeiten, Inhalte zu annotieren und zu kommentieren, sind mit „Lesespuren“ auch in analogen Medien gegeben. Doch dass solche „Lesespuren“ interaktiv erfolgen, auf verschiedenen Levels veröffentlicht werden können und grundsätzlich auf Dauer verfügbar sind, ist gegenüber bisherigen Formen der Kommentierung von Dokumenten anders und neu. Zugleich geht das Potenzial digitaler Interaktivität weit darüber hinaus, wenn die dafür verfügbaren Werkzeuge des Web 2.0 genutzt werden.

(3) „Enhanced Publications“ sind „processable“, und diese Eigenschaft bietet ebenfalls große Mehrwerte, die sich weit über herkömmliche Analyseverfahren hinaus ausschöpfen und nutzen lassen. Denn im Unterschied zu analogen Medien lassen sich digitale Medien maschinell verarbeiten; das bedeutet, dass „Enhanced Publications“ einschließlich ihrer Komponenten (Bilder, (Retro-)Digitalisate, Forschungsdaten etc.) mit Methoden des Data- und Text-Minings analysiert, ausgewertet und durchsucht werden können.

(4) „Enhanced Publications“ sind „linkable“, was für digitale Publikationen von großem Vorteil ist. Konnten analoge Publikationen bisher maximal über ihre elektronischen Metadaten miteinander in Beziehung gesetzt werden, so eröffnen digitale Publikationen die zusätzliche Möglichkeit, mit den Metadaten auch die Inhalte selbst zur Verfügung zu stellen und zu kontextualisieren. Dieses Potenzial an Verlinkungsoptionen, das bis zu semantischen Vernetzungen reicht, stellt ebenfalls eine neue Qualität dar, die zugleich eine deutlich bessere Vernetzung ermöglicht, als dies bisher gegeben ist.

10 Degkwitz, A., Digitale Sammlungen – Vision eines Neubeginns. – In: Bibliothek, Forschung und Praxis 38 (2014) 3, S. 411 – 416.

Die genannten Kriterien lassen eine neue Qualität erkennen, die für digitale Medien kennzeichnend ist. Damit einher gehen Veränderungen der Publikationskette, für die weiterhin das Peer-Reviewing eine zentrale Rolle spielt, deren Prozesse jedoch nicht mehr aufeinander aufbauen, sondern sehr viel stärker als bisher in einem Vernetzungskontext stehen. Auch wird und kann die am Ende stehende Publikation auf verschiedenen Wegen der Wissenschaftskommunikation distribuiert werden, was insbesondere in Verbindung mit dem Open-Access-Publizieren zu einer erhöhten Wahrnehmbarkeit von Forschungsergebnissen führt. Diese deutlich gesteigerten Möglichkeiten der Verbreitung und Vermittlung von Inhalten können und sollten nun noch dadurch erweitert werden, dass mit „Enhanced Publications“ Inhalte zur Verfügung stehen, die maschinell verarbeitet und somit nachgenutzt werden können. Dies gehört entweder zu den künftigen Aufgaben von Verlagen, die die zur Veröffentlichung vorgesehenen Inhalte entsprechend aufbereiten sollten, oder zum Servicespektrum von Bibliotheken, die die Aufbereitung von Veröffentlichungen nach ihrem Erscheinen vollziehen. Entscheidende Voraussetzungen dafür sind rechtliche Rahmenbedingungen, die die Aufbereitung und Verarbeitung publizierter Forschungsergebnisse erlauben. Doch an dieser Stelle stehen Bibliotheken und Verlage noch ganz am Anfang. Mit der Verarbeitbarkeit von Inhalten sind insbesondere Bibliotheken in der Situation, (1) entsprechende Werkzeuge zur Verfügung zu stellen, (2) für Metadatenbeschreibung, Textstrukturierung und Referenzierung von Inhalten und Objekten geeignete Dienste bereitzustellen und (3) im Rahmen ihres Sammlungs- und Versorgungsauftrages die langfristige Verfügbarkeit zu gewährleisten. Das sind schon jetzt und auf jeden Fall künftig zentrale Erwartungen an digitale Bibliotheken und ihre Sammlungen.

#### *4. Zusammenfassung*

Die Transformation des Wissens und der Wandel des Publizierens stehen in einem engen Zusammenhang, der in den digitalen Medien sichtbar wird. Dies zeigt sich an der zunehmenden Vielfalt an Wissensrepräsentationen und damit an Medien- und Publikationsformen, die weniger Texte als vielmehr Daten veröffentlichen und damit Forschungsergebnisse stärker visualisieren als sie - wie bisher - zu verbalisieren. Doch über das Aggregationspotenzial digitaler Publikationen hinaus verändert sich auch deren Nachnutzung durch neue Möglichkeiten der Kontextualisierung, Verarbeitung und Verlinkung. Geht es dabei in der Tat nur um digitale Infrastrukturen, die wissenschaftliches Arbeiten auf eine neue Grundlage stellen? Und geht es dabei nur um Publikationen, die die digitalen Potenziale für die Veröffentlichung von Forschungsergebnissen nutzen? In jeder

Technik steckt „Intelligenz“, die eingesetzt wird, um menschliche Intelligenz zu entlasten oder zu unterstützen. So kann es kaum anders sein, dass sich menschliche und technische Intelligenz wechselseitig beeinflussen und dieser wechselseitige Einfluss auch für Infrastrukturen, Medien und Wissen gilt. Mit Blick auf die eingangs angesprochenen Forschungsformen, die sich gegenseitig nicht ausschließen, sondern vielmehr ergänzen, eröffnen sich neue Wissensräume, ohne dass „alte“ Wissensräume deshalb geschlossen werden – vielmehr weiten und erweitern sich unsere Wissensräume in ihrer digitalen Transformation.

Zeiten des Wandels sind von dem Spagat geprägt, dass diejenigen Akteure, die für den Wandel plädieren, oftmals solche sind, für die alles Gegenwärtige viel zu fest an Vergangenen klebt, um sich im Regelwerk der Routine bewegen zu können. Dem gegenüber stehen Akteure, die sich dem Wandel entziehen, weil sie an oft selbst gesetzten Regeln festhalten möchten, um sich nicht Künftigem auszusetzen, das keine Regeln kennt. Das führt zu der Allerweltsfrage: Was gewinnen wir mit der digitalen Transformation und was geht uns möglicher Weise verloren? Doch diese Frage ist vergeblich, weil ihre Beantwortung sich nicht aus dem Wandel ergibt und der Wandel selbst sie nicht beantworten kann. Zudem ist sie falsch gestellt, weil ein ultimatives „entweder-oder“ gar nicht gegeben ist. Wissenschaftliches Arbeiten kennt disziplinspezifische Forschungsformen und Methoden, die sich nicht gegenseitig ausschließen müssen, sondern sich vielmehr gegenseitig ergänzen. Mit dem erweiterten Materialspektrum, das sich vor allem den Geisteswissenschaften mit digitalen Verfahren erschließt, geht deshalb nichts verloren, sondern es stellen sich alte Fragen neu, und neue Fragen lösen alte Fragen ab. Zugleich treten Bibliotheken und Wissenschaft in eine Kooperation, die eine neue interdisziplinäre Qualität erreicht und die für die Weiterentwicklung beider Partner äußerst Gewinn bringend ist. Zu adressieren ist allerdings, dass bewusst werden sollte, was man mit digitalen Infrastrukturen und Verfahren tatsächlich macht. Doch dergleichen fordert schon die Neugier am eigenen Handeln aller Akteure, die in Forschung und Infrastruktur „digital“ unterwegs sind.